NPS ARCHIVE 1966 MCCULLOUGH, D.

> AN APPLICATION OF QUEUEING THEORY TO THE OPERATION OF REPLENISHMENT AT SEA

DAVID UNDERWOOD MCCULLOUGH

LIBRARY NAVAL POSTGRADUATE SCHC^ MONTEREY, CALIF. 93940

.



AN APPLICATION OF QUEUEING THEORY TO THE OPERATION OF REPLENISHMENT AT SEA

by

David Underwood Mc Cullough Lieutenant, United States Navy

Submitted in partial fulfillment for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

UNITED STATES NAVAL POSTGRADUATE SCHOOL October 1966

ACCULLOUGH, D.

ABSTRACT

A simplified model of the Navy's underway replenishment operation is investigated by using a multi-stage cyclic queueing model to approximate the process. A computer simulation of a more realistic model is performed to generate data which is compared with the cyclic queue results. The cases for M = 2 and N = 10, 15, 20 with various service rates are considered, where M is the number of ships in the underway replenishment group, and N is the size of the combatant force being replenished. The following two measures of effectiveness are considered:

- T, the time required for the underway replenishment group to process N combatant force ships.
- (2) WC, the number of combatant force ship-hours required for the replenishment.

The cyclic queue and computer simulation results agree reasonably well for a balanced system. However, in general, the cyclic queue results provide only an upper bound for the computer simulation results.

The simulation model demonstrates for the balanced system that those combinations of initial starting sequences which result in the lowest values of T also result in the highest values for WC. The converse, that is, low values of WC resulted in high values of T, was also shown to hold for the simulation model.

TABLE OF CONTENTS

and int, the second

Section		Page
1.	Introduction to the Queueing Problem	9
2.	Series of Queues	11
3. 7	The Queueing System	13
4. 1	Representation of the Queueing Process	15
5.	Two Queues in Series	19
6. 1	Formulation of the Problem	22
7. /	Analytic Approximation to a Solution	27
8. 8	Simulation Approach to a Solution	36
9. (Conclusion and Recommendations	51
10. /	Acknowledgement	52
11. H	Bibliography	53

the second second

and the second se

and the second second

---- 0

11 P.L.L.C. 11

LIST OF ILLUSTRATIONS

Figur	e	Page
1.	Typical Initial Arrangement for Underway Replenishment of a Combatant Force by a URG.	24
2.	Initial Arrangement Assumed for First Analysis.	24
3.	Logic Flow Diagram for Computer Simulation Model.	38
4.	Plot Showing Variation of T _{avg} and T _{var} with Replication Size.	47



Table		Page
I	Analytic Approximation Results When All URG Ships Have the Same Mean Service Rate ($u = 2$) with $M = 2_{\circ}$	33
II	Analytic Approximation Results When One of the URG Ships has Double the Service Rate of the Other ($u = 4$, $u_2 = 2$) with $M = 2$, 1	34
III	Analytic Approximation Results When One of the URG Ships has Triple the Service Rate of the Other $(u_1 = 6, u_2 = 2)$ with $M = 2, 1$	35
IV	Computer Simulation Results with $M = 2$.	49

 \sim

1. Introduction to the Queueing Problem.

A common phenomenon occurring in everyday life is that of "queueing" or waiting in a line. Queues (waiting lines) form, for example, at bus stops, supermarket counters, and ticket booths. Queues are also found in the Navy, such as on the mess deck where men wait to receive their food and eat, in storerooms where the parts wait to be used, and in underway replenishment operations where combatant ships wait to receive food, fuel, and ammunition from the ships of an underway replenishment group. It is this latter operation which we will investigate.

The theory of queues¹ is concerned with the development of mathematical models for examining the behavior of systems that provide service for randomly arising demands. In a wider sense it deals with the investigation of the probability law of different processes occurring in connection with mass servicing in cases where random fluctuations occur. The practical aim in investigating a queueing system is to improve the system by changing it in some way.

The problems responsible for the theory of queues seem to have arisen in the telephone industry and the work of A. K. Erlang on telephone traffic problems, carried out as early as 1908, constitutes the first major contribution to the theory of queues.

Queueing occurs when demands for service are greater than availability of service, and it becomes necessary to postpone the demands by a system of queueing or marshalling. While queueing is a natural process, marshalling is an attempt to do something about it. Queueing is

Queueing theory is the label preferred by the British and Indians who have been particularly active in the mathematical analysis of the process.

essentially a temporary phenomenon, as otherwise the queue would grow indefinitely in size. Thus, one of the first problems in any queueing situation is to determine whether or not the demand will outstrip the service mechanism. When this does not happen, so that the queue is continually returning to zero size, then the situation is said to be "in equilibrium", or in statistical terminology, the process is "stationary". When the demand equals the service capabilities the situation becomes very delicately balanced; a slight reduction in demand yields a sensible queueing system that can attain equilibrium, while a slight increase results in an ever lengthing queue.

The theory has been applied to a number of problems seemingly diverse in nature, but most deal with the following type situation: A "customer" arrives at the "counter" and demands service. If the server is busy with another customer, the newly arrived customer must wait until the server is free. In the meantime, other customers may arrive at the service. If customers arrive when the server is busy, they must form a queue or waiting line until service is available. Here we use the terms "customer" and "counter" in a generic sense. In the present paper they will correspond to combatant ship and replenishment ship respectively.

2. Series of Queues.

The majority of the published studies of queues deal with the type of problem in which one service operation is performed on each customer, although one or more service channels in parallel may be involved. A more general type of problem, which has not been treated so extensively, is one in which customers are served at each of a number of counters arranged in series; every customer begins service at a first counter and continues through the system, receiving attention at all others, until he is discharged at a last counter. Series of queues or queues in tandem occur in a variety of applications. One example is that of customers in a store who must first be waited on by sales clerks and then, after being served by these clerks, must then be served by wrappers or cashiers. Another example is the U.S. Navy's at sea replenishment operation in which combatant ships must receive fuel, food, and ammunition from the underway replenishment group (URG) which is composed of ships that individually can supply only fuel, only food, only stores, or only ammunition. Combatant ships which require all three of these services thus become involved in a tandem queueing system with the URG ships acting as servers.

Most of the analytical work in queues with a number of facilities in series has been restricted to Poisson arrivals and exponential service times. The studies have almost entirely concentrated on deriving steady-state solutions.

An important question to determine in studying tandem queues is the distribution of output from one channel which then comprises the input into a subsequent channel.

Reich [10] has shown that for interarrival and service periods having normalized chi-squared distribution with four degrees of freedom and for single-channel queue, the departure epochs do not constitute a normalized chi-squared distribution with four degrees of freedom. Hence, it is not generally reasonable to expect the outputs to match the inputs for a general interarrival pattern, even in the case of the steady state.

3. The Queueing System.

In order to describe a given queueing system, it is necessary to specify [5] the following components of the system: (1) the input process, (2) the queue discipline, and (3) the service mechanism.

The input process is usually expressed by a probability law governing the arrival of customers at the counter where service is provided. Suppose the customers arrive at the counter at times t_1 , t_2 , ..., t_n ($t_1 < t_2 < \ldots < t_n$). Let $T_n = t_{n+1} - t_n$ denote the difference between the time of arrival of the (n + 1)st and the nth customers. The input process is given by the probability law governing the sequence of arrival times $\{T_n\}$. However, the queueing system considered in this paper does not have a probabilistic input. But rather the customers (combatant ships) are initially distributed throughout the system of counters (URG ships) and then the queueing process begins.

The queue discipline is the rule or moral code determining the manner in which customers form a queue and the manner in which they behave while waiting. In this paper we assume that the queue discipline can be expressed as "first come, first served". (Other possibilities are to select a customer at random with respect to order of arrival, or to take the last customer to arrive rather than the first, or to service customers in batches, or to select a customer according to some priority rule, or and so forth.)

The service mechanism can be described as follows: Let the random variable E_n denote the time required to serve the nth customer; hence the probability law governing the sequence of service times $\{E_n\}$ expresses the service mechanism of the queueing system.

It is natural to assume that the successive service times $E_1, E_2, ..., E_n, ...$ are statistically independent of one another and of the sequence of interarrival times $\{T_n\}$ and that they have the same distribution function B(E), $E \in (0, \infty)$. There are many distribution functions of interest, one of which is

B(E) = 1 -
$$e^{-uE}$$
, for E≥0
= 0, for E<0.

This is called the exponential distribution and it is widely used in the investigation of queueing systems. In the above, u is the expected or mean service time. The exponential distribution of service times will be used in this paper since observational data are lacking and it often provides a good fit to real life situations. 4. Representation of the Queueing Process.

The stochastic processes arising in the theory of queues are in general non-Markovian and it is only for systems of the type having Poisson arrivals, exponential service, and multi-servers in parallel that the associated processes are Markovian. Hence it is necessary to consider various ways of representing queueing systems so that their stochastic properties can be ascertained.

In order to investigate the stochastic properties of a particular queueing system, it is necessary to formulate a mathematical representation (model) which is based on (1) the input process, (2) the queue discipline, and (3) the service mechanism which characterize the queueing system. A queueing system is describable at any fixed point in time in terms of a state vector; such a vector must contain sufficient information for calculation of the joint distribution function of its components at all subsequent points of time. The simplest example of a state vector consists of the single element n, the number in the system, but in more complicated cases we may require such components as the elapsed time since the last arrival, or the elapsed time since each customer in service commenced service.

It is usually of importance to discover if the joint distribution function depends on the time since the state vector had a known value. In many cases after sufficient time the function becomes independent of time and of the known state. Such systems are said to reach a "steady state", described by the limiting form of the distribution function. Other factors of interest are the distribution function for customer waiting time, mean and variance of the number in the system, the number served, the probability of waiting longer

than a given period, the expected number of idle servers (for multiserver queues), and the expected idle time for each server.

There are several main techniques which have been used in the representation and analysis of a queueing system. One of the most often used is the differential-difference equation method. This technique is applicable to queueing systems which have Poisson inputs, exponential service, and servers in parallel. Queues of this type are commonly called Poisson queues and have Markovian properties. The method, as the name implies, relies on the formation of a series of differential equations connecting the probabilities of the states and substates with their first derivatives. Specifically, if we let X(t) denote the number of customers in the queue at time t, then $\{X(t), t \ge 0\}$ is a Markov process with a denumerable number of states and its stochastic properties can be derived from the Kolmogorov differential equations representing the process. Now if $P(t) = (p_{ij}(t))$ denotes the matrix of transition probabilities associated with the process $\{X(t), t \ge 0\}$, then P(t) satisfies the system of Kolmogorov equations

$$\frac{d}{dt} P(t) = P(t)A(t),$$

where $A(t) = (a_{ij}(t))$ is the matrix of infinitesimal transition probabilities and P(0) = I, the identity matrix. In order to solve these equations, it is necessary to specify the functions $a_{ij}(t)$, which in the study of queueing systems will generally be functions of time, and the parameters characterizing the interarrival-time and service-time distribution functions. The theory of queues based on the Kolmogorov equations can be divided into two areas or parts, one of which is called non-equilibrium theory and the other equilibrium theory. In the equilib-

rium theory interest centers on the probabilities $\mathbb{W}_{x} = \lim_{t \to \infty} P(X(t) = x)$, i.e., one is interested in finding the limiting or stationary probability distribution $\{\mathbb{W}_{x}\}$ for the states x. These probabilities, if they exist, are obtained by solving the Kolmogorov equations when the time derivative is put equal to zero, since in the steady-state the probabilities do not change with time and the time variable may be dropped. In the non-equilibrium theory interest centers on the probabilities that at time t the queue is of length x, i.e., P(X(t) = x). These probabilities are obtained by solving the system of Kolmogorov equations. This technique will be used in Section 7 for deriving the steady-state distribution of a special type queueing system.

Other techniques of interest are the Embedded Markov Chain Method [6], Lindley's Integral Equation Representation [8], Integrodifferential Equation Method of Takacs [12], and the Multi-dimensional state-space approach [4].

The various techniques mentioned above are described as being analytic as opposed to simulation and Monte Carlo methods. Instead of making a mathematical analysis of the properties of the queueing system [2], it may be advisable to examine the process by reconstructing its behavior using service times, arrival times, etc., derived from random numbers. This approach is particularly useful when the process is so complicated that mathematical solution is likely to be difficult or impossible, and especially when the behavior is required under very special and clearly defined conditions and no mathematical solution is immediately available. It may then happen that empirical sampling is likely to lead to an answer in a reasonable time, whereas the effort needed to produce a mathematical solution may be

difficult to gauge.

The simplest procedure is to use random numbers to construct a direct realization of the queueing process corresponding as closely as possible to the real system. In order to obtain more precise conclusions for a given amount of effort, it may, however, be profitable to modify the process that is sampled. There is no universally accepted terminology for this method, but the term Monte Carlo is often reserved for a procedure in which the process sampled has been modified to increase precision. For example, if the quantity of interest can be broken into components some of which can be calculated theoretically and some not, precision will usually be increased by sampling only the components that cannot be found theoretically. The term simulation is used when the process sampled is a close model of the real system. An advantage of simulation over a Monte Carlo method is that the detailed results give a direct qualitative impression of what the system should look like under the conditions postulated. Both the analytic approach and the simulation techniques will be used to investigate the system considered in this paper.

5. Two Queues in Series.

For an infinite input with Poisson distribution with parameter λ to the first of two channels in series, with exponential service times having parameter u_1 and u_2 respectively, at each channel (a customer must go through both service channels), Jackson [3] has derived the Kolmogorov equations involving $P(n_1, n_2, t)$, the probability that there are n_1 units in the first stage of the system (including service) and n_2 in the second at time t. The steady-state solution is obtained by setting the derivative portion of these equations equal to zero and solving for $P(n_1, n_2)$. The steady-state solution is

$$P(n_1, n_2) = \left(\frac{\lambda}{u_1}\right)^{n_1} \left(\frac{\lambda}{u_2}\right)^{n_2} P(0, 0)$$
(5.1)

The value of P(0,0) is determined to be

 $P(0,0)=\left(1-\frac{\lambda}{u_1}\right)\left(1-\frac{\lambda}{u_2}\right).$

assuming that $\frac{\lambda}{u_1}$ and $\frac{\lambda}{u_2}$ 1.

An interesting result is expressed by equation 5.1. It may be shown in the case of a single stage queue having Poisson (parameter λ) inputs and exponential service times (parameter u) that the steady-state probability of having n units in the system is given by

$$P(n) = \left(\frac{\lambda}{u}\right)^n P(o)$$

Thus equation 5.1 shows that in a two stage Poisson queueing system the stages act independently of one another in the steady-state condition.

Nelson [9] has derived the waiting time distribution for this type queue, $P(\sum_{j=1}^{2} X_{j} \leq T)$, as a function of the waiting time distribution at the individual stages $P(X_{j} \leq T)$, j = 1,2. He assumes

$$P(x_{j} \leq T) = K_{j} e^{c_{j}T}$$

for the probability of waiting longer than time T at the jth stage where

$$c^{\beta} = n^{\beta}(1 - \frac{n^{\beta}}{\gamma})$$

KJ= b(xg>o)

The culmulative probability distribution for the total waiting time through both stages is shown to be

$$P\left(\sum_{j=1}^{2} x_{j} \leq T\right) = I - A_{j} \approx e^{c_{j}T}$$

where

$$A_{\partial x} = K_{\partial} \left(\frac{\pi}{\pi} \frac{1 - \kappa_{i} c_{\partial}}{c_{\partial} - c_{i}} \right), \quad \mathcal{F}^{=1,2}.$$

and $c_i \neq c_j$ for $i \neq j$. In arriving at this result Nelson makes use of the results of Burke's investigation of the output of a queueing system. Burke [1] has shown that the steady-state output of a queue with s channels in parallel, with Poisson input and parameter λ and the same exponential service time distribution with parameter u for each channel is itself Poisson distributed with the same parameter λ as that of the input distribution. He also showed the independence of the inter-departure-interval random variables and the state of the system at the end of the interval.

6. Formulation of the Problem.

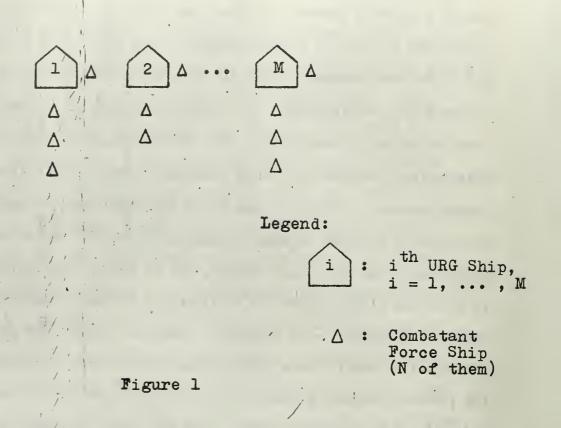
The prime desideratum in replenishing a combatant force at sea by means of an underway replenishment group is to replenish it as quickly as possible while observing the necessary rules of safety. The time required to conduct the replenishment is of basic importance due to the extreme vulnerability of the combatant force during replenishment operations. In addition, the combatant force has little operational value while being replenished; also both combatant and replenishment forces are usually operating on a tight and inflexible schedule. For these reasons, the replenishment at sea operation must be conducted with upmost dispatch consistent with safety. One intuitive solution to the problem of minimizing the force replenishment time is to minimize the individual ship replenishment times. Other solutions involve finding optimal combinations of initial conditions and sequence for replenishment. Thus in attempting to minimize the individual ship replenishment time, one must investigate the behavior of the combatant ships during the underway replenishing process. This fact makes necessary a quantitative analysis of the underway replenishing process, with a view towards more efficient utilization of the underway replenishment group. It thus seems natural to apply the theory of queues in the analysis. Let us consider the main aspects of the situation under the three headings introduced by Kendall to characterize a queueing process.

As a preliminary step, consider the pattern of events prior to the actual replenishment operation. The underway replenishment group (URG), composed of refueling ships, ammunition ships, and reprovisioning ships

is in one of the prescribed underway replenishment formations and steaming on a suitable course. The combatant force is in formation astern of the URG. The ships of the combatant force awaiting to be replenished have been assigned a sequence for replenishment based upon individual ship requirements. For example, one ship of the combatant force might first receive fuel, then ammunition, then food, and finally spare parts. Another ship would probably be assigned a different sequence: however, if ships happen to be assigned identical sequences they receive services in order of seniority with the senior most ship goint first, the next senior second, and so forth. The basic problem to be solved at this point can be stated as follows: What sequences should be assigned to the respective combatant force ships so as to minimize the overall force replenishment time? This is an example of the job-shop sequencing problem [11]. It is a problem for further study and will not be considered here. When the signal is given to begin the replenishment, the combatant ships break formation and go to the first URG ship in their respective replenishment sequences where they either commence replenishment or join a queue. Thus one can visualize a series queue in which the customers are combatant force ships proceeding in accordance with their respective sequences and the servers are URG ships.

First, there is the input process. In our context this is essentially the process by which the combatant force ships break formation and feed into the URG. Thus if the combatant force is composed of N ships and the URG is composed of M ships we may visualize M queues and associated waiting lines arranged initially in a manner similar to the diagram of Figure 1. In our first analysis we shall assume that the input process is not as shown in Figure 1, but rather

Typical Initial Arrangement for Underway Replenishment of a Combatant Force by a URG



Initial Arrangement Assumed for First Analysis

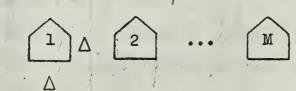


Figure 2

 Δ

Δ

Legend:



1

 i^{th} URG Ship, $i = 1, \dots, M$

∆ : Combatant Force Ship (N of them) as shown in Figure 2, with all N ships queueing up to the first URG ship. This assumption is made to make the analysis tractable and represents a particular case of the input process of Figure 1. Second, there is the queue discipline. In the first analysis we shall require that the combatant force ships join the URG at the 1St URG ship and proceed through the sequence of URG ships, leaving the system only when all M stages have been completed. First come, first served discipline will be used at each stage. Third, there is the service mechanism, which is given by the frequency distribution of service times. We will make the simplifying assumption that the service rate for the ith URG ship is exponential with parameter u. This assumption is made for three reasons: (1) observational data are lacking, (2) it is desired to make the analysis tractable, and (3) the exponential distribution often gives a good fit to real life.

Before proceeding with the analysis it should be noted that one fundamental difference between the general queueing problem and the problem at hand is that we are not concerned with the limiting case, that is an infinitely long queueing process. The queue associated with this problem has a finite bound N as a datum of the problem and must be included in the analysis. In order to do this we will impose the artificial restriction that upon completing the replenishment at the Mth URG ship, a combatant ship rejoins the system again at the 1st URG ship. This technique allows us to have in effect an infinite queueing system.

In one sense, the problem of minimizing force replenishment time is mathematically trivial because it is finite. In principle, one could deal with this problem simply by enumerating all possible solu-

tions, evaluating each according to whatever criterion is relevant and selecting the best. This approach, of course, is feasible only for problems of small magnitude. For example, a typical replenishment operation involving M = 4 URG ships and N = 20 combatant force ships has 1771 different possible sequences and even with the aid of highspeed computing equipment an enumeration of all solutions is not feasible. In general, the feasibility of straight forward enumeration is perhaps even more remote for large scale problems than the simple example noted here. Because the problem is finite, even though possibly enormous, a "method of solution" cannot merely be a way of arriving at a correct answer, but must be a technique where by this answer is obtained at low computational cost, at least relative to the cost of simple enumeration and evaluation. Thus, the object is nor merely to solve, but to solve efficiently and the analytic approach to solving the problem represents an attempt at achieving an efficient solution.

7. Analytic Approximation to a Solution.

As an analytic approximation to a solution of the problem of minimizing combatant force replenishment time we consider a system with M sequential stages in a loop; each stage acts as a single server. The system serves N units, each of which goes through all stages in sucession and continuously repeats the process. The technique which we present for analyzing a system of this type was originally formulated by Koenigsburg [7].

A state is indicated by (n_1, n_2, \dots, n_M) where $\sum_{k=1}^{M} n_i = N$ and n_i indicates the number of units being served or waiting for service at the ith stage. A waiting unit has completed service at the (i - 1)st stage. The probability of being in a state (n_1, n_2, \dots, n_M) is written $P(n_1, n_2, \dots, n_M)$.

Transitions between states occur when a unit enters or leaves a stage. We have assumed that the service rate in each stage is a random variable which can be described by a mean value u_i and a distribution which is exponential, i.e.,

ui A t = probability that a unit being serviced has completed in the time interval between t and t + At (u = mean service rate).

The probability that the system is in state (n_1, n_2, \dots, n_M) at t = t + Δ t is expressed as the sum of probability terms. Since we assume that transitions only occur from nearby states, they can occur when the system was in the states $(n_1, n_2, \dots, n_{i-1} + 1, n_i - 1, \dots, n_M)$ and $(n_1, n_2, \dots, n_i + 1, n_{i+1} - 1, \dots, n_M)$, at the time t.

$$P(n_{1},...,n_{m})(t+\Delta t) = P(n_{1},...,n_{m})(t) \Big[1 - (\mu_{1}+\mu_{2}+...+\mu_{m})\Delta t \Big] + P(n_{1}+1,n_{2}-1,n_{3},...,n_{m})(t) \mu_{1}\Delta t + P(n_{1},n_{2}+1,n_{3}-1,...,n_{m})(t) \mu_{2}\Delta t + ... + P(n_{1}-1,n_{2},...,n_{m}+1)(t) \mu_{m}\Delta t$$

This can be written as

$$P(n_{1}, n_{2}, ..., n_{M}) (\pm + \Delta t) - P(n_{1}, n_{2}, ..., n_{M}) (\pm)$$

$$\Delta t$$

$$= -\sum_{i=1}^{M} U_{i} P(n_{1}, n_{2}, ..., n_{M})$$

$$+ U_{i} P(n_{i+1}, n_{2}^{-1}, ..., n_{M})$$

$$+ ...,$$

$$+ U_{M} P(n_{i-1}, n_{2}, ..., n_{M}^{+1})$$

The left-hand side is the defined derivative of $P(n_1, n_2, \dots, n_M)$ so we can write

$$\frac{d}{dt} P(n_1, ..., n_m) = -\sum_{\lambda=1}^{m} u_{\lambda} P(n_1, ..., n_m) + \sum_{\lambda=1}^{m} u_{\lambda} P(n_1, n_2, ..., n_m)$$

subject to the following restrictions:

(a) \$\sum_{i}^{M}\$ \$n_{i} = N\$
(b) if \$n_{i} = 0\$, then the \$u_{i}\$ term = 0 (term does not enter into equations)

(c) if
$$n_{i+1} - 1 < 0$$
, then the u_i term = 0.

(d) the Mth stage is linked back to the first stage; i.e., transitions occur from (n₁ - 1, n₂, ..., n_M + 1) to (n₁, n₂, ..., n_M).

Since this technique is used to arrive at only an approximate analytic solution, we are interested in the steady-state condition, i.e.,

$$\frac{d}{dt} P(n_1, n_2, ..., n_M) = 0$$
(7.1)

Thus there are $\frac{(N + M - 1)!}{(M - 1)!}$ equations in the same number of unknowns

(this is the number of ways of putting N things into M boxes any number to a box).

It may be verified by direct substitution in equation (7.1) that this system of equations has solutions of the form

$$P(n_1, n_2, ..., n_M) = \frac{u_1^{n_2} u_3^{n_3} ... u_M^{n_M}}{u_2^{n_2} u_3^{n_3} ... u_M^{n_M}} P(n_1, 0, 0, ..., 0)$$

where

$$P(N,0,0,\dots,0) = \frac{1}{\sum_{\substack{X \neq 0 \\ N}} \frac{1}{X_{2}^{n_{1}} x_{2}^{n_{2}} \cdots x_{m}^{n_{m}}}}$$

$$= \left[\sum_{\substack{A = 1 \\ \partial^{z_{1}} \\ \lambda = 1 \\ \lambda = 1 \\ \lambda \neq \partial}}^{N} \frac{1}{\sum_{\substack{X \neq 0 \\ N}} \frac{1}{\sum_{\substack{X \neq 0 \\ X \neq 0}}}} \right]^{-1} = \frac{1}{Z_{m}^{N}}$$

and

$$X_{\lambda} = \frac{u_{1}}{u_{\lambda}}, \lambda = 1, \dots, M$$

Now given a particular URG ship, say i, it is working whenever $n_i \ge 1$. The fraction of the time it does not work is given by

$$D_{\lambda} = \sum_{n_{\lambda}=N} P(n_{1}, n_{2}, ..., n_{\lambda-1}, 0, n_{\lambda+1}, ..., n_{M})$$

$$= P(N, 0, ..., 0) \sum_{n_{\lambda}=1} X_{1}^{n_{1}} X_{2}^{n_{2}} ... X_{M}^{n_{M}}$$

$$= \frac{1}{2} \frac{1}{2}$$

where the i in ${}_{i}Z_{M}^{N}$ -1 indicates X_{i} is omitted from the summation.

The mean number of units at the ith stage is given by

$$\overline{n}_{A} = \sum n_{A} P(n_{1}, ..., n_{A}, ..., n_{M}) = \frac{1}{Z_{M}^{N}} \sum_{AII \text{ Partitions}} n_{AII} x_{AII}^{n_{I}} ... x_{M}^{n_{M}}$$
All Partitions of N

$$= \frac{X_{i}}{Z_{m}^{N}} \frac{d}{dX_{i}} Z_{m}^{N}$$

The mean number of units awaiting service at the ith stage is given by

$$\overline{\omega}_{\lambda} = \sum_{\substack{\text{Allpointdivis} \\ o \in N}} (n_{\lambda} - i) P(n_{1}, \dots, n_{M})$$

$$\xrightarrow{\text{Allpointdivis} \\ o \in N}$$

$$= \frac{1}{R_{M}^{N}} (\overline{n}_{\lambda} \overline{z}_{M}^{N} - \overline{z}_{M}^{N} + \overline{z}_{M-1}^{N})$$

$$= \overline{n}_{\lambda} - 1 + D_{\lambda}^{\lambda}$$

The mean cycle time (the time to complete service at all stages) is

$$T = \sum_{A=1}^{\infty} \frac{1}{u_A} + \sum_{A=1}^{\infty} \frac{\overline{w}_A}{u_A}$$

From this we see that the mean cycle time is independent of the order in which the URG ships are arranged.

In the special case where all URG ships have the same mean service time, i.e., where $u_1 = u_2 = \dots = u_M$ and $X_1 = X_2 = \dots = X_M = 1$, we have

The mean number of units at the jth stage is by definition

$$= \frac{1}{Z_{m}^{N}} \sum_{\substack{n_{n=1} \\ n_{n=1} \\ m}}^{M} \frac{n_{n} x_{n}^{n_{n}} \sum_{\substack{n_{n} \\ n_{n} = 1}}^{n_{n}} \frac{n_{n} x_{n}^{n_{n}} \sum_{\substack{n_{n} \\ n_{n} = 1}}^{n_{n}} \frac{n_{n} x_{n}^{n_{n}} \sum_{\substack{n_{n} \\ n_{n} = 1}}^{n_{n}} \frac{(n + m - n_{n-2})!}{(m - 2)! (n - n_{n})!}$$

$$= \frac{1}{Z_{m}^{N}} \frac{(m + m - 1)!}{(m + 1)!}$$

This result is apparent from the statement of the problem. The number of units waiting at the jth stage is, by definition

$$\overline{\omega}_{j} = \frac{1}{Z_{m}^{N}} \sum_{\substack{A \mid i \in N \mid A \mid n \mid n \\ A \mid i \in N \mid A \mid n \mid n \\ a \in N \\ n_{A} \ge i \\ = \overline{n}_{j} - 1 + \frac{M - 1}{N + M - 1}$$
$$= \frac{N(N - 1)}{M(N + M - 1)}$$

The mean cycle time is therefore

 $T = \frac{M}{u} + \frac{M\overline{w}}{u}, \quad \overline{w} = \overline{w}_{\partial}, \quad \partial^{=1,\dots,M}$

Tables I, II, and III were compiled using the formulas developed in this section. Table I shows the results when all ships of the URG have the same mean service rate (u = 2) with M = 2. Table II shows the results when one of the URG ships has double the service rate ($u_1 = 4$, $u_2 = 2$) of the other with M = 2 and Table III shows the results when one of the URG ships has triple ($u_1 = 6$, $u_2 = 2$) the service rate of the other. It is apparent from these data that a significant reduction in mean cycle time is obtained by doubling the service rate of one of the URG ships, but no such reduction is realized over the double service rate when the service rate of one of the ships is tripled. This latter fact suggests that the system's performance is dominated by the slowest server when a large unbalance exist between service rates.

TABLE I

Analytic Approximation Results When All URG Ships Have the Same Mean Service Rate (u=2)with M = 2

Legend:

- N = number of units in the system
- \overline{n} = mean number of units in each stage
- \overline{w} = mean number of units waiting at each stage
- T = time required for a unit to complete M stages (in hours)

N	n	w	Т
10	5	4.09	5.09
15	7.5	6.56	7.56
20	10.0	9.05	10.05

TABLE II

Analytic Approximation Results When One of URG Ships Has Double the Service Rate of the Other (u = 4, u = 2) with M = 2. 1 2

Legend:

N = number of units in the system
n
= mean number of units at stage i
w
= mean number of units waiting at stage i
T = time required for a unit to complete M st

Γ	Ξ	time	required	for	а	unit	to	complete M stages	
		(in 1	nours)						

N	i	n _i	w _i	Т
10	1	0,985	0.000	
10	2	9,001	8.002	4,75
15				
15	1 2	0,999	0.000	7,24
	2	13,335	12,555	/ , 24
20	1	0.999	0.000	
	2	19,000	18.000	9.75

TABLE III

Analytic Approximation Results When One of the URG Ships Has Triple the Service Rate of the Other (u = 6, u = 2) with M = 2.

Legend:

N = number of units in the system

n = mean number of units at stage i

w = mean number of units waiting at stage i

T = time required for a unit to complete M stages (in hours)

N	i	n_i	w	Т
10	1	0.499	0.000	1
	2	9.400	8.500	4.71
15	1	0.500	0.000	
	2	14.300	13.500	7.22
20	1	0.500	0.000	
	2	19,200	18.200	9.71

8. Simulation Approach to a Solution

The preceeding section provides an example of an approximate analytic approach to the solution of the problem of minimizing force replenishment time. It should be clear that this effort has led to relatively meager results. Intuitively, these results represent only bounds for a more accurate solution. Moreover, a survey of the literature indicates that presently there is little sign of a major breakthrough in achieving a more useful analytic solution to the problem. In these circumstances, attention will be directed to techniques which are less than wholly analytic, and in particular to computational simulation--or perhaps "computational experimentation" is more appropriate phrase if we seek general conclusions rather than specific solutions. The purpose of experimentation, strictly speaking, cannot be the determination of optimal sequences. One can compare the effectiveness of alternate proposed sequences. Or, one can seek more generally to identify, measure, and eventually to predict a variety of relations between various sequences and their consequences.

We have developed a computer model of a two-stage series queue that has initial conditions similar to the initial conditions of the underway replenishment operation and which simulates the movement of combatant ships through the URG. The technique used in developing this simple model may be used for developing the model of a system with any number of stages.

The model constructed is an incremental time step computer simulation of the interactions between the combatant force and the URG during an underway replenishment operation. The purpose of the model is to test the various sequences and to make comparisons with the re-

sults derived from the analytic approximation previously described. The following assumptions were made in constructing the model:

- Service at stage j, i.e., at ship j of the URG, is exponential with parameter u (one server at each stage), j = 1,2.
- (2) Customers (combatant ships) are initially distributed among the various stages with I customers at stage j, j = 1,2.
- (3) Upon completing service at their initial stage, customers join the queue at the other stage for service.
 Upon completing service at this stage they leave the system.
- (4) Service at each stage is first-come, first-served.
- (5) Customers pass through the system in the same cyclic order.

The logical structure of the model is outlined in Figure 3. Since the service times are exponential, i.e., the probability of a customer being served has completed in the time interval between t and t + Δ t is $\underline{u} \cdot \underline{\Delta}$ t, the computational procedures are relatively easy and fit in very nicely with the incremental time step concept. The flow diagram demonstrates the pattern which may be used for developing a model which simulates any number of stages.

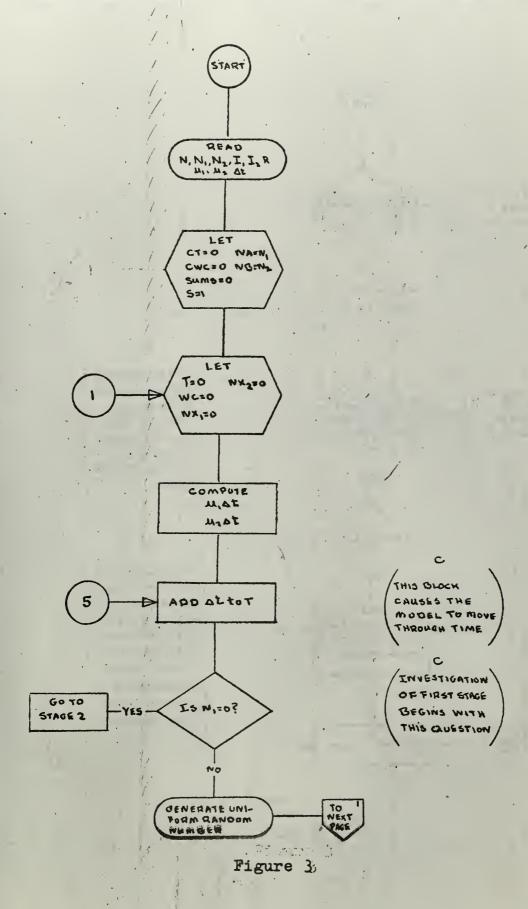
In order to use the model the following information must be provided:

 a. N, the number of combatant ships to receive services from the URG.

Logic Flow Diagram for Computer Simulation Model

List of Symbols Used in the Flow Diagram:

CT	Cumulative total time (used for computing Tave)
CWC	Cumulative total combatant ship-hours in system
Δt	Incremental time step
I	Initial number of combatant ships at stage i, i = 1,2
N	Number of ships in the combatant force
NA,NB	Temporary storage locations
Ni	Number of combatant ships at stage i, i = 1,2
NX.	Number of combatant ships served at i th stage, i = 1,2
R	Replication size
S	Number of replications completed
SUMS	Sum of squares of the T's
T	Time to process N combatant ships thru system
Tavg	Average over the replications of T
Tvar	Variance over the replications of T
ui	Service rate of the i th URG ship, i = 1,2
wē	Combatant ship-hours spent in the system
WCavg.	Average over the replications of combatant ship-hours spent in the system



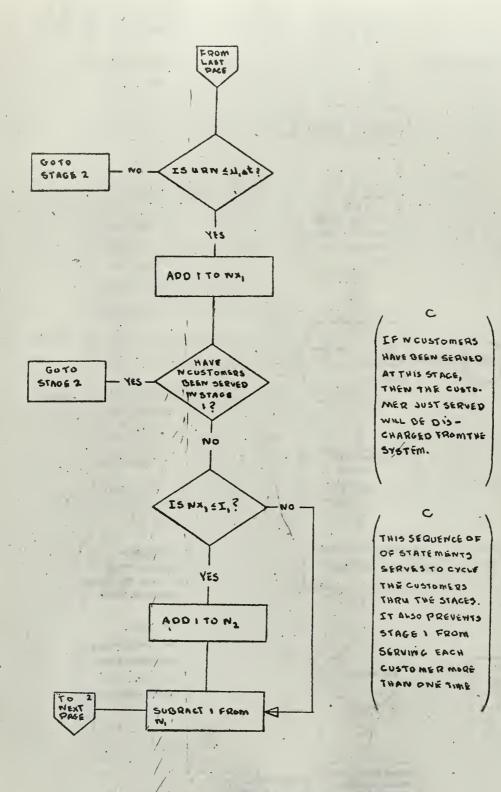
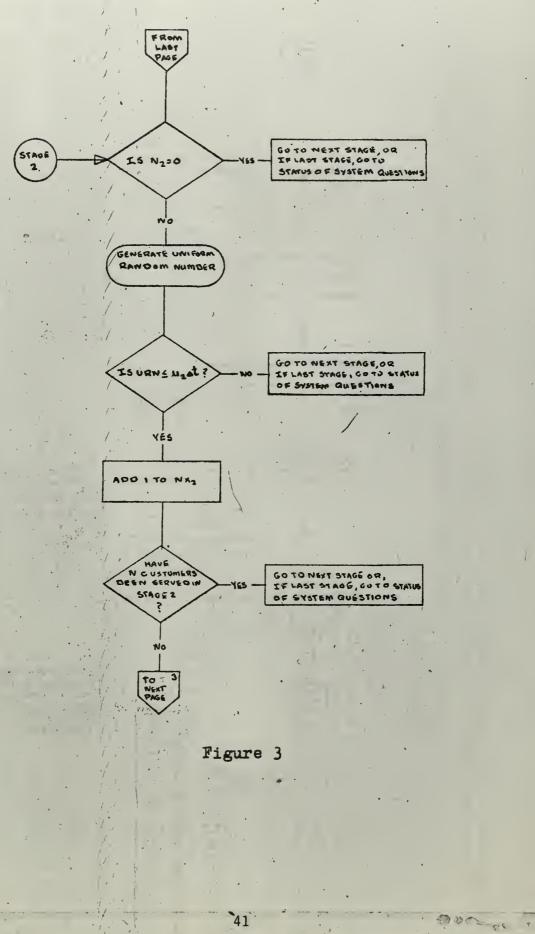
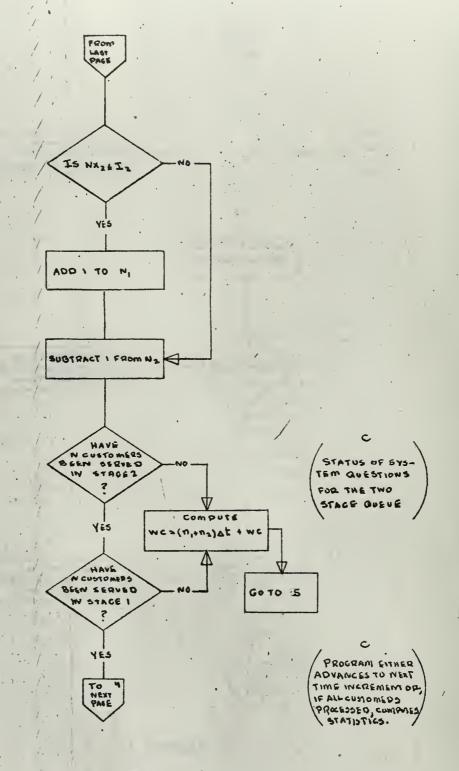


Figure 3.

Logic Flow Diagram for Computer Simulation Model (cont.)

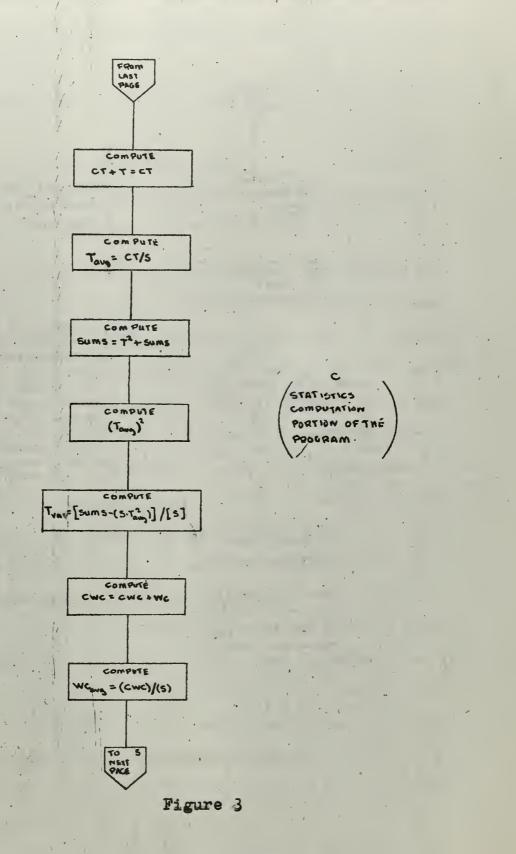






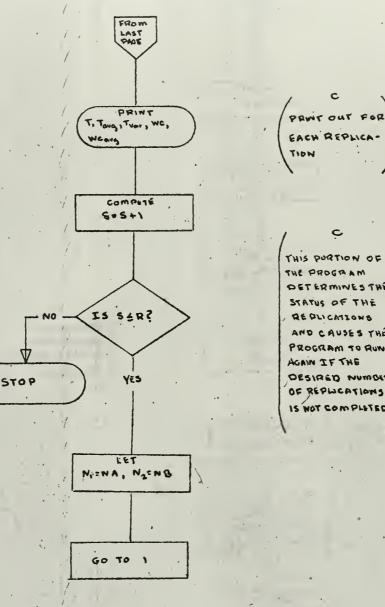
Logic Flow Diagram for Computer Simulation Model (cont.)

11



Dorwer "

Logic Flow Diagram for Computer Simulation Model (cont.)



IDN THIS PORTION OF THE PROGRAM

c

11

DETERMINESTHE STATUS OF THE REPLICATIONS AND CAUSES THE PROGRAM TO RUN AGAIN IF THE DESIRED NUMBER OF REPLICATIONS IS NOT COMPLETED

Figure 3

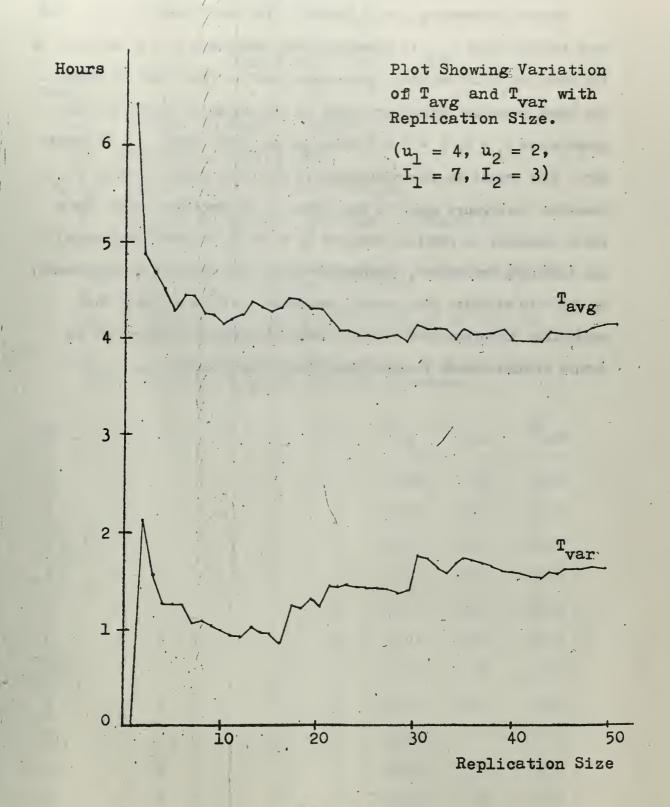
- b. u_j, the service rate (in ships per hour) of the jth URG ship, j = 1,2.
- c. I, the number of combatant ships initially at stage j,
 j = 1,2.
- d. \triangle t, the incremental time step.
- e. R, the number of times a particular set of I, u, and N are j, j, in to be simulated, i.e., the replication size.

Although a model of this type may be used to generate data for various measures of effectiveness applicable to the system, we have restricted ourselves to examining only two, namely, (1) the mean time, T_{avg} , for the system to process N customers, and (2) the total number of combatant ship-hours required for the replenishment. Others might be URG ship-hours required for the replenishment, mean idle time for the ith URG ship, etc. The first of the two measures considered may be compared with the mean cycle time derived from the analytic approximation in order to test how well the two approaches agree. The second measure is not obtainable from the analytic approximation and represents the actual number of combatant ship hours involved in the replenishment. Implicit in the second measure is the following assumption: Combatant ships which have finished their replenishment may proceed without having to wait for the remainder of the combatant force. The numerical difference between the second measure and the quantity Tava times N represents the number of ship-hours lost if the entire combatant force waited until the Nth ship completed its replenishment before proceeding.

Two measures of performance of the simulation itself have been included as a check on the statistical significance of the models out-

put. The first of these is a running average over the replications of T, where T is the time obtained, from one replication, for the system to process N units. The average over all the replications of T is T_{avg} . The other measure of performance is the running variance over the replications of T. Figure 4 is a typical graphical plot of these running variables versus replication size. It is apparent from this figure that the model is performing satisfactorily, since the running variables seem to be converging to definite values.

Sixteen cases were investigated by computer simulation in order to obtain data which could be used for comparison with the analytic approximation results and for comparison between cases. The results of the cases investigated are tabulated in Table IV; also included in Table IV is a listing of the analytic approximation results from Tables I, II, and III which are of interest. It is interesting to note for the cases where both $u_1 = u_2$ and $I_1 = I_2$ that the analytic approximation and computer simulation results agree very closely. However, this agreement fails to hold for the other cases and the analytic approximation provides only an upper bound for the simulation The analytic results do describe a characteristic of the results. system which is confirmed by simulation data. This characteristic of the system has to do with the decrease of T_{avg} as the mean service rate of one URG ship is first double, and finally triple, the mean service rate of the other URG ship. By making the service rate of the first URG ship double that of the second a significant decrease in T_{avg} is noted; however, when the service rate of the first URG ship is triple that of the second the decrease in $T_{av\sigma}$ with respect to the double service rate case is negligible.





47

1.5

Another interesting result occurs in the cases where $u_1 = u_2$. The data indicate that T is lowest for the case where $I_1 = I_2$ and this is the case for which combatant ship-hours spent in the system is highest. The total combatant ship-hours spent in the system is lowest for the cases where $I_1 = N$, $I_2 = 0$ and these are the cases where T is greatest. This result is also noticeable in the other cases where $u_1 \neq u_2$. Combatant ship-hours spent in the system is an important factor for a force commander to consider whenever he must deploy individual ships, say screening destroyers, immediately after they complete replenishment. In order to minimize this factor, sequences could be assigned that would take advantage of the result just mentioned. A portion of any future studies should further investigate this result.

TABLE IV

Computer Simulation Results with M = 2

Legend:

N	= number of ships in combatant force
М	= number of ships in URG
u _i	= mean service rate of i th URG ship
I _i	<pre>= number of combatant force ships initially at URG ship i</pre>
Tavg	<pre>= average, over fifty replications, of the time for URG to process N combatant force ships</pre>
^T var	<pre>= variance, over fifty replications, of the time for URG to process N combatant force ships</pre>
WCav	= average, over fifty replications, of the number of combatant ship-hours spent in the system

Case	N	u 1	u 2	1 1	1 ₂	Tavg	Tvar	WCavg
1	10	2	2	5	5	5.096	1.136	33.69
2	10	2	2	7	3	5.182	1.963	33.78
3	10	2	2	10	0	5,432	1.877	30.92
4	10	4	2	10	0	4.307	1,715	23,79
5	10	4	2	7	3	4.153	1.653	24.67
6	10	4	2	5	5	4.207	1.664	25.11
7	10	4	2	3	7	4.199	1.763	24.28
8	10	4	2	0	10	4.241	1.680	23.51
9	10	6	2	10	0	4.149	1.139	21.83
10	10	6	2	7	3	4.183	1.042	23.92
11	10	6	2	5	5	4.036	1,190	22.90
12	10	6	2	3	3	4.078	0.907	21.47
13	10	6	2	0	0	4.064	0.781	22.55

TABLE IV (continued)

)

Case	Ν	u ₁	u ₂	I ₁	¹ 2	Tavg	T var	WCavg
14	20	2	2	20	0	10.650	5.774	114.08
15	20	2	2	15	5	10,259	4.445	127.16
16	20	2	2	10	10	10.054	4.713	134.08

Analytic Approximation Results of Interest (see Tables I-III)

10	2	2		5.09
10	4	2		4.75
10	6	2		4.71
20	2	2	,	10.05

í.

•

9. Conclusion and Recommendations.

An investigation of the underway replenishment operation has been conducted by examining two simplified models of the actual process. The object of the investigation was to determine methods for reducing the overall time spent in replenishing a combatant force. The first model considered, an analytic approximation to the actual process, suggested that force replenishment time may be reduced by increasing the service rate of the URG ships. This increase should not occur for just one ship of the URG, but must be realized for all of them; since continuously increasing the service rate of only one of the URG vessels results in a decreasing marginal gain in overall force replenishment time. The second model considered was a computer simulation of the replenishment process. This model confirmed the results of the analytic approximation, but illustrated that the analytic model, in general, serves only to provide an upper bound for force replenishment times. However, in the case of a balanced URG, i.e., all URG ships have the same service rate, the analytic approximation provides a good extimate of force replenishment time.

The investigation conducted in this paper could be extended by developing models which more closely approximate the underway replenishment process. Future models could include URG ships having multiservers and combatant force replenishment sequences other than cyclic.) In view of the difficulty encountered in achieving even a simple analytic model, the most useful future models will probably be computer simulations. Although the exponential distribution of service times probably gives a good approximation for use in investigating the under-

way replenishment operation, other distributions, such as the Erlang, should be considered in future studies.

10. Acknowledgement

The writer is deeply indebted to Professor Stephen M. Pollock for his guidance and encouragement through the course of this investigation.

BIBLIOGRAPHY

- 1. Burke, Paul J. The output of a queueing system. Operations Research, vol. 4, 1956: 699-704.
- 2. Cox, D.R. and Smith, W.L. Queues. Methuen, 1961.
- 3. Jackson, R.R.P. Queueing systems with phase type service. Operational Research Quart., vol. 5, 1954: 109-120.
- 4. Keilson, J. and Kooharian, A. A general model for queues. (Unpublished) from the Applied Research Laboratory, Sylvania Electronic Systems, Waltham, Mass.
- 5. Kendall, D.G. Some problems in the theory of queues. J. Roy. Statist. Soc., Ser. B. vol. 13, no. 2, 1951: 151-185.
- 6. Kendall D.G. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. Ann. Math. Statist., vol. 24, 1953: 338-354.
- Koenigsberg, E. Cyclic queues. <u>Operational Research Quart.</u>, vol. 8, no. 2, March-April, 1958: 22-35.
- 8. Lindley, D.V. The theory of queues with a single server. Proc. Cambridge Phil. Soc., vol. 48, 1952: 277-289.
- 9. Nelson, R.T. Waiting-time distributions for application to a series of service centers. Operations Research, vol. 6, 1958: 856-862.
- 10. Reich, E. Waiting-times when queues are in tandem. Ann. Math. Statist., vol. 28, no. 3, Sept., 1957: 768-773.
- 11. Sisson, R.L. Methods of sequencing in job shops a review. Operations Research, vol. 7, no. 1, 1959: 10-29.
- 12. Takacs, L. Investigation of waiting time problems by reduction to Markov processes. Acta. Math. Acad. Sci. Hung., vol. 6, 1955: 101-109.

INITIAL DISTRIBUTION LIST

		No.	Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314		20
2.	Library U.S. Naval Postgraduate School Monterey, California 93940		2
3.	Prof. Stephen Pollock (Thesis Advisor) Department of Operations Analysis (Code 55 Pd) U.S. Naval Postgraduate School Monterey, California 93940		1
4.	Prof. S. Parry (Code 55 Py) Department of Operations Analysis U.S. Naval Postgraduate School Monterey, California 93940		1
5.	Lt. David Underwood Mc Cullough, USN Box 146 Fayette, Missouri		1

Unclassified		
Security Classification	-	
DOCUMENT CO	ONTROL DATA - R&D	
(Security classification of title, body of abstract and index		
1. ORIGINATIN G ACTIVITY (Corporate author)	2 <i>a</i> . 1	Unclassified
U.S. Naval Postgraduate School Monterey, California	2 5. 0	sroup N/A
3. REPORT TITLE		**/ **
An Application of Queucing Theory	ory to the Operation At Sea	n of Replenishment
 DESCRIPTIVE NOTES (Type of report and inclusive dates) N/A 		
5. AUTHOR(S) (Last name, first name, initial)		
MC CULLOUGH, David U.		
6. REPORT DATE	78. TOTAL NO. OF PAGES	75. NO. OF REFS
October 1966	53	12
a. CONTRACT OR GRANT NO.	94. ORIGINATOR'S REPORT	NUMBER(S)
N/A		
b. PROJECT NO.	N/A	
N/A		(Annually second s
c.	this report)	(Any other numbers that may be assigned
d.	N/A	
10. AVAILABILITY/LIMITATION NOTICES		
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY	ACTIVITY
group to pro	yclic queueing mode more realistic mode e cyclic queue result service rates are replenishment group d. The following two required for the ur press N combatant for er of combatant for enishment. ion results agree to he cyclic queue results. rates for the balance es which result in WC. The converse,	el to approximate the el is performed to gen- its. The cases for considered, where M is , and N is the size of wo measures of effective- nderway replenishment orce ships. rce ship-hours required reasonably well for a bal- sults provide only an up- ted system that these com- the lowest values of T , that is, low values of
DD 150RM 1473		Unclassified

Unclassified Security Classification

Unclassified Security Classification

14.				LINI	KA	LINK	(8	LIN	кс
KEY WORDS			ROLE	WT	ROLE	WT	ROLE	WT	
									1
underwa	ay replenishment								
	<i>xy</i> representation								
series	queue				1		1	1	
series	queue simulation								
							1	-	
			-						
			-				1		
						1			
			_		_			0	
							2		
									_

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (corporate author) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking ia to be in accordance with appropriate security regulations.

2b. GROUP: Automatic downgrading is apecified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report ss day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7 a. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8e. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9e. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (either by the originator or by the sponsor), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard ststements such as:

- (1) "Qualified requesters may obtain copies of this report from DDC."
- (2) "Foreign announcement and dissemination of this report by DDC is not authorized."
- (3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
- (4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
- (5) "All distribution of this report is controlled. Qualified DDC users shall request through

1 22

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (paying for) the research and development. Include sddress.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as (TS), (S), (C), or (U).

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.



